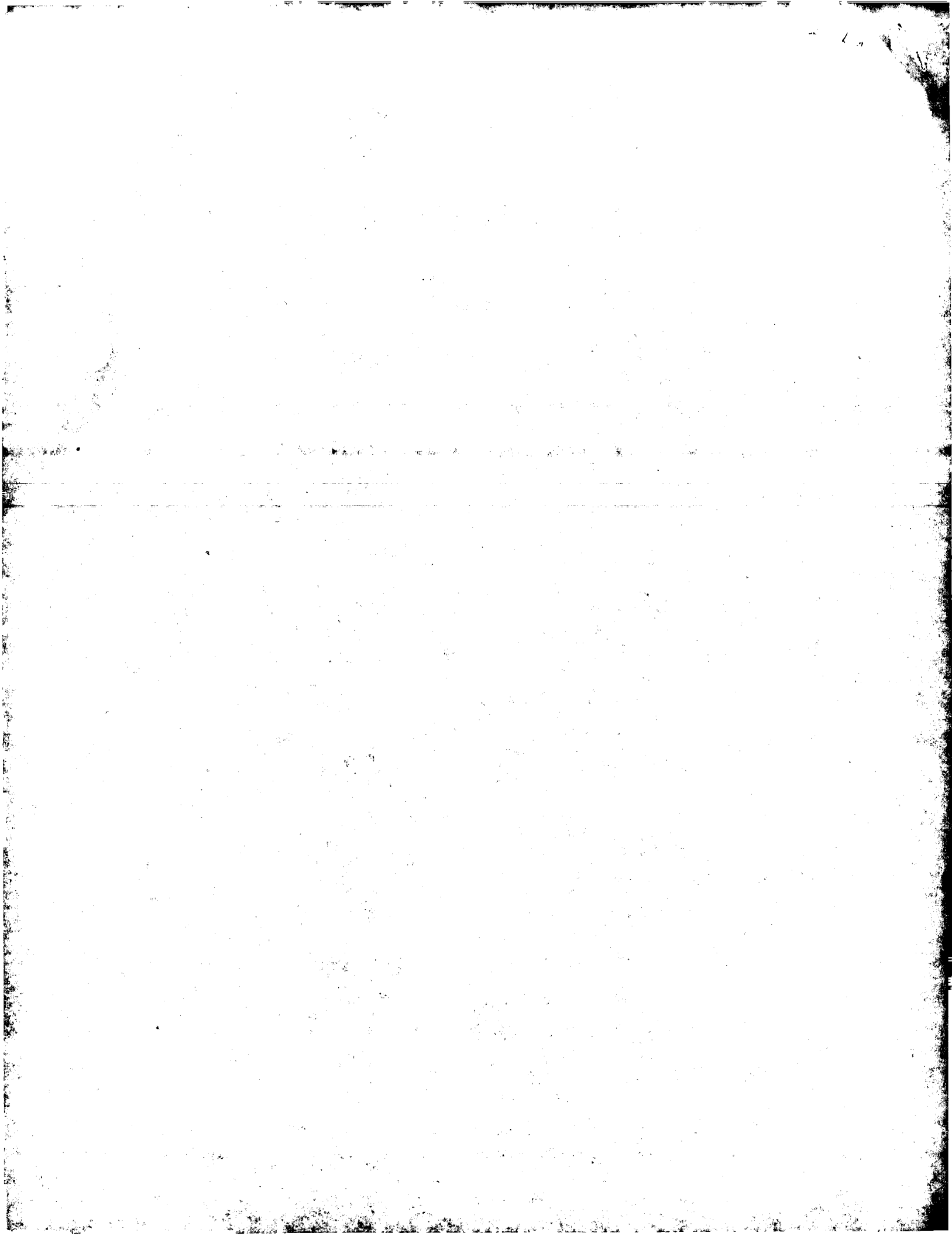


!!NA_SEQUENCE 1.0
LOCUS HSM800235 1433 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922).
ACCESSION AL049450
VERSION AL049450.1 GI:4500236
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1433)
AUTHORS Koshner, R., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp586B1922) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
Location/Qualifiers
1..1433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp586B1922"
/db_xref="taxon:9606"
/clone="DKFZp586B1922"
/tissue_type="uterus"
/clone_id="586 (synonym: hute1). Vector psport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
polyA_signal
polyA_site 1386..1391
BASE COUNT 414 a 321 c 331 g 367 t
ORIGIN
AL049450 Length: 1433 July 21, 2003 13:18 Type: N Check: 450 ..
1 TAGGAAGAAG CGCTCAACAC AGAACACCT AAGTATCCCT ACTCCCTTTC
51 AAGGACTAAT TCCCTTCAGA GATTTCCTT GCTTCAGTCC TTCAGAGACC
101 TCCCGAAGAT GCTAGAGATC GTTTAAGAG GCTGGAGAT GCCCTTGGG
151 AAGGAATGT TGTAGCTGG CAACTACTCA TCCCAAGAG GCGGTGGCAA
201 TAGCTCCCTG GGCAGCAGAC ATCATGATTT TCCCTCCCTC TGGGAACGAG
251 AGTCCCTCC GTTCTGCTT TGAATCCCTT AGAGACAAAG AACTTCATTC
301 TCTCCCATAC CCGCGTGGGA CTGCTGGGA GGAGAGGGG TGGAGGGTT
351 AGACTTCTAC CCCCATTTGG AAGTACAGAA AGAAGACTTC CTGTTAAAAA
401 GCTGGCTCT GAATTTCAA AAGTCAAAC AAAAATCTCT AAGCACTCTC
451 CTATTAAAGA GGAACCTGT GGTTCCTTAT CTGAAACTGT TTGTAACGT
501 GAATTGAGA GCCAAGAAAC CCCAGAAAG CCCGGTCTT CAGTGGACAC
551 CCCACCAAGA CTCTCCACTC CCCAAAGGG ACCCAGACAC CATCCCAAGG
601 AGAAAGCTT CTCAAGTAG ATAGAGATT TGGCGTACTT TTCACACACA
651 GAAATGATT TGTGTGTTG GCACACAGCT CCCCATTCAC CGTTACATT
701 ACGGGAATCC TCTCAAGA AGAGAGAC TTAGCAAGT AAGCATAGA

751 GAACACTTGC TCTTATACCC TAGTGTGGC GGTCAAGCTA ACAAGTGTGA
801 AATGCTCTT GGCATTTTAA AAAAGTGCA ATCAATTAAG CAGAGTCTG
851 TCAAGATGA GTAGTTAC AGCCAGAGAC AAGACTGTG CAGCATTCG
901 AATAGATGG AATTACAGCA AAATGTCTC AATGATTTTG CCGCTTACA
951 ACATGGAG ATGTGTTGC CAGTAAGTTG CTCATCACAA GAGCAGCAGA
1001 CTGGGGGTG TAATCTCCG CAATGTGCA GCCCTCTGAA AGAAGGTTT
1051 TCTGTCTGT GAAATGCATA GAATATACT TTGCATGCA CGACTGTTC
1101 TGCAATGAT ATTGTGAA ATCTGGAGG GTGTCTTTG GGTGTTCTCA
1151 GGGGCAATG GTATTTTGG GGTGGGAG CCAGCTTGG GTGGGAATT
1201 TTACCTGGG CTCGCTCTT TTAATATAT AAATTTTAT CTGATATCT
1251 ATGCTCTGT CTGGGGGCA GGAGAACTT GCCAAAGACC AACAGTCTTA
1301 CTTATCTTA CTATCTCA CAAAGTCTT AAATGTGA GAGTTACTT
1351 GATTCAGT AGCCATTTG TTGTCATAT ATTAAATA AATGCTTAC
1401 AACTATTTT TCAACAAAA AAAAAAAAA AAA



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2003, 13:29:30 ; Search time 3 Seconds
(without alignments)
4.177 Million cell updates/sec

Title: HSN800235

Perfect score: 1433
Sequence: 1 TAGGAAGAGCGCTCAACAC.....AACAAAAAAAAAAAAAAAAA 1433

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 segs, 4372 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Database : US10054935.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query length | ID | Description |
|------------|-------|-------|--------------|----|-----------------------------------|
| 1 | 380.2 | 26.5 | 4372 | 1 | US-10-054-935-1 Sequence 1, Appli |

ALIGNMENTS

RESULT 1
US-10-054-935-1
; Sequence 1, Application US/10054935
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
; FILE REFERENCE: 16U 107 R1
; CURRENT APPLICATION NUMBER: US/10/054,935
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1922)
; OTHER INFORMATION:
US-10-054-935-1

Query Match 26.5%; Score 380.2; DB 1; Length 4372;

Best local Similarity 95.6%; Pred. No. 0;
Matches 391; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Db 1045 GCGGAGTGGAAAGGACATATAAGGAATCCCTTTGGAGTACAGAAAGAAAGACTC 1104
QY 391 CTGTTAAAAAGCTGCTCCTGAATTTTCAAAAAGTCAAAAACAAAACCTCTAAGACACTC 450
Db 1105 CTGTTAAAAAGCTGCTCCTGAATTTTCAAAAAGTCAAAAACAAAACCTCTAAGACACTC 1164
QY 451 CTATTAAAGAGGAACTGCTGCTTATCTGAAGCTGTTTGAAGTGAATGAGGA 510
Db 1165 CTATTAAAGAGGAACTGCTGCTTATCTGAAGCTGTTTGAAGTGAATGAGGA 1224
QY 511 GCGAAGAAACCCAGAAAAGCCGCTTCAGTGAACCCACCAAGACTCTGCACATC 570
Db 1225 GCGAAGAAACCCAGAAAAGCCGCTTCAGTGAACCCACCAAGACTCTGCACATC 1284
QY 571 CCAAAAAGGAGCCAGCAACCCATCCCAAGAGAAAGCCCTTCAAGTAGATAGAAGATT 630
Db 1285 CCAAAAAGGAGCCAGCAACCCATCCCAAGAGAAAGCCCTTCAAGTAGATAGAAGATT 1344
QY 631 TGCCGTACCTTTCCACACAGAAATGTAATTTGTGCTTGGACACAGCTCCCATCAG 690
Db 1345 TGCCGTACCTTTCCACACAGAAATGTAATTTGTGCTTGGACACAGCTCCCATCAG 1404
QY 691 GGTACCAATTACGGGAATCTCTCCAAAGAGAGAGACTGTAGCAG 739
Db 1405 GGTACCAATTACGGGAATCTCTCCAAAGAGAGAGACTGTAGCAG 1453

Search completed: July 21, 2003, 13:29:34
Job time : 3 secs


```

!INA_SEQUENCE 1.0
LOCUS      HSM800235          1433 bp      mRNA      linear      PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA: cDNA DKFZp586B1922 (from clone DKFZp586B1922).
ACCESSION  AL049450
VERSION     AL049450.1  GI:4500236
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1433)
AUTHORS     Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE       Direct Submission
JOURNAL     Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     This clone (DKFZp586B1922) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES
    source          1..1433
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="RZPD:DKFZp586B1922"
                    /db_xref="taxon:9606"
                    /clone="DKFZp586B1922"
                    /tissue_type="uterus"
                    /clone_lib="586 (synonym: hute1). Vector pSport1; host
                    DH10B; sites NotI + SalI/MluI"
                    /dev_stage="adult"
    polyA_signal    1386..1391
    polyA_site      1412..1413
    BASE COUNT      414 a      321 c      331 g      367 t
    ORIGIN
AL049450 Length: 1433 July 21, 2003 13:18 Type: N Check: 450 ..
1  TAGGAGAGAG CGCTCAACAC AGACACCTT AAGTATCCT ACTCCCTTTC
51  AAGGACTTAA TCCCTTCAG GATTTCCTT GCTTCAGTC TTCAGAGCC
101  TCCCGAGAT GCTAGAGTC GTTAAAGAG GCTGGAAGT GCCCGTTGG
151  AAGGAACGT TGTAGCTGG CAATACTTCA TCCCAAGGAG GCGGTGCCAA
201  TAGCTCCCTG GGCAGCAGAC ATCATGAAAT TTCCTCCCTC TGGGAACGAC
251  AGGTCCCTCC GTTCTCGTT TGAATCCCT AGAGACAAG AACCTTCATTC
301  TCTCCCATAC CCGGTGGGA CTGCTGGGA GGAGAGGGG TGAAGGGTT
351  AGACTTCTAC CCCATTGG AGTACAGAA AGAAGACTC CTGTTAAAA
401  GCTGCTCCT GAATTTCAA AAGTCAAA AAAAACTCT AAGCACTCTC
451  CTATTAAAG GGAACCTGT GGTTCCTTA CTGAACCTG TTGTAACGT
501  GAATTGAGA GCCAAGAAC CCCAGAAAG CCCCGTCTT CAGTGACAC
551  CCCACCAAG CTCCTCACT CCCAAAGG ACCCAGCAC CATCCCAAG
601  AGAAAGCCT CTCAAGTAG ATGAAGATT TCCCGTACT TTCCACACA
651  GAATGTATT TGTGTGTTG GCACCAAGC CCCCATCAC GGTACCAAT
701  ACGGAATCC TCTCAAGA AGAGAGGAC TGTAGCAAG AAGCATAGA

```

```

751  GAACACTTGC TCTTATACC TAGTGTGGC GGTCAAGCA ACAAGTGTG
801  AATGCTCTT GCAATTTTA AAAAGTGCA ATCAATTAAG CAGAGTCTG
851  TCAAGATGA GTAAGTTAA AGCCAGAGC AGACACTGTG CAGCATTTG
901  AATAGATGG AATTACGCA AATGTGCTC AATGATTTG CCTGCTTCA
951  ACACGTGGAG ATGTGTTGC CAGTAAGTTG CTCATCACA GAGCACCAGA
1001  CTTGGGGGTG TAATCTCCG CAACCTGCAT GCCCTCTGAA AGAAGGTTT
1051  TCTGTCTGT GAAATGCATA GAACATTAAT TTGCCATGCA GCACTGTTCC
1101  TGCATTTGAT ATTGTGAA ATCTGGAGG GTGTCCTTG GGTGTTCTCA
1151  GGGGCCAATG GTAATTTTG GGTGGGGAG CCAGCTTGG GTGGGAATT
1201  TTCACCTGGG CTTCCGCTT TTAACATAT AACAATTAT CTGATATCT
1251  ATGTCCCTGT CTGGGGGCA GGAGAACTT GCCAAGACC AACAGTCTTA
1301  CTTATCTTA CTATCTTCA CAAAGTCTT AAATGTGAA GAGTTTACTT
1351  GGATTGCAGT AGCCATTTG TTGTTCAAT ATTAAATTA AATGCTCTAC
1401  AACTATTTT TCAAAACAAA AAAAAAAAA AAA

```


11NA_SEQUENCE 1.0
LOCUS HSM800235 1433 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA: cDNA DKFZp586B1922 (from clone DKFZp586B1922).
ACCESSION AL049450
VERSION AL049450.1 GI:4500236
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1433)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp586B1922) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
location/Qualifiers
1.1433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp586B1922"
/db_xref="taxon:9606"
/clone="DKFZp586B1922"
/tissue_type="uterus"
/clone_lib="586 (synonym: hute1). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1386..1391
polya_signal
polya_site 1412
BASE COUNT 414 a 321 c 331 g 367 t
ORIGIN
AL049450 Length: 1433 July 21, 2003 13:18 Type: N Check: 450 ..
1 TAGAAGAG CGCTCAAC AGAACCCCT AAGTATCCCT ACGCCCTTC
51 AAGACTAAA TCCCTTCAGA GATTTCCTTT GCTTCAGTCC TTCAGAGACC
101 TCCCGAAGAT GCTAGAGATC GTTTAAGAAG GCTGGAAGAT GGCCGTTGGG
151 AAGGAAGTGT TGCTAGCTGG CAATACTTCA TCCCAAGAGG GCGGTGGCAA
201 TAGCTCCCTG GCGACAGAC ATCATGAGAT TTCTCCCTC TGGAACCCAG
251 AGGTCCCTCC GTTTCGTGTT TGAATCCCTT AGAGACAAG AACCTCATTC
301 TCTCCCATAC CCGCTGGGA CTGCTGGGA GAGAGGGGG TGGAGGGTT
351 AGACTTCTAC CCCCATTTGG AAGTACAGA AGAAGACTC CTGTTAAAA
401 GCTGGCTCCT GAATTTTCAA AAGTCAAAAC AAAAATCCTT AAGCATCTC
451 CTATTAAAGA GGAACCTGT GGTCTCTTAT CTGAACTGT TTGTAACGT
501 GAATTGAGGA GCCAAGAAC CCCAGAAAAG CCCGGTCTT CAGTGACAC
551 CCCACCAAGA CTCTCCACTC CCCAAAGGG ACCCAGCAC CATCCCAAG
601 AGAAGCCTT CTCAGTGAG ATAGAAGATT TGCGTACCT TTCACACACA
651 GAATGTATT TGTGCTTG GCACCAAGCT CCCCATCAC GGTACCATT
701 ACGGAATCC TCTCAAGA AGAGAGAGAC TGTAACAAGT AAGCATAGA

751 GAACACTTGC TCTTATACC TAGTGTGC GGTCAAGCTA ACAAGTGTA
801 AATGACCTTT GGCATTTTAA AAAAGTGA ATCAATTAAG CAGAGTCTG
851 TCAAGAAATGA GTAAGTTAC AGCCAGAGAC AGACACTGTG CAGGCATTC
901 AATAGATAG AATTACAGCA AATGTGCTC AATGATTTTG CCTGCTTACA
951 ACACCTGGAG ATGTGTTTC CAGTAAGTTC CTCATCACAA GAGACACAGA
1001 CTTGGGGGTG TAATCTCCG CAACCTTCAT GCCCTGTGAA AGAAGGTTT
1051 TCTGTGCTGT GAAATGCATA GAACATTAAT TTGCCATGCA CGACTGTTCC
1101 TGCAATTGAT ATTGTGGA ATCTGGAGG GTGGCTTTG GGTGTTCTCA
1151 GGGGCCAATG GTAATTTTGG GGTGGGGAG CCAGCTTGGG GTGGGGAAT
1201 TTCACTGGG CTTCCGCTCT TTAATATAT AAACATTTAT CTGTATATCT
1251 ATGTCCCTGT CTGGGGGCA GAGGAATCT GCCAAGACC AACGTCTTA
1301 CTTATCTTA CTATCTTCA CAAAGTCTT AAATGTGAA GAGTTTACTT
1351 GGATTGCAGT AGCCATTTGG TTGTCATAT ATTAAATAA AATGCTTAC
1401 AACTATTTT TCAACAAAA AAAAAAAAA AAA

